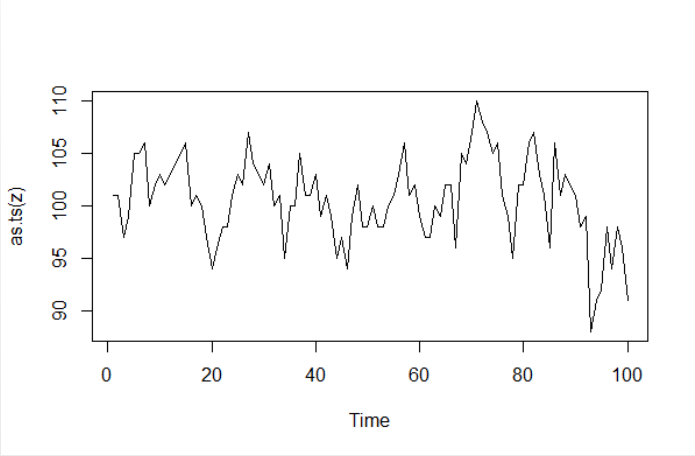
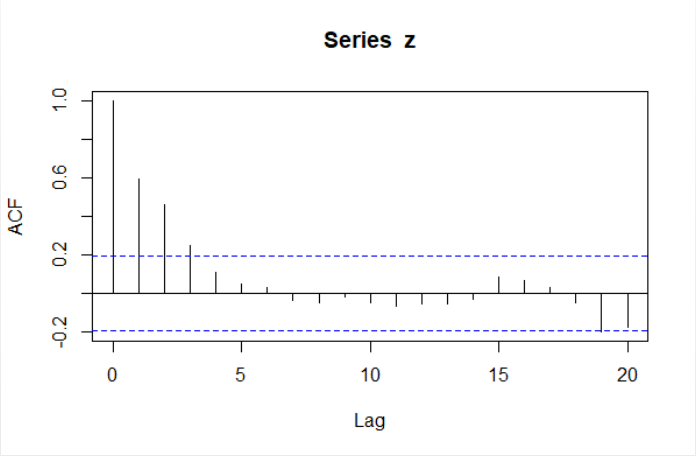
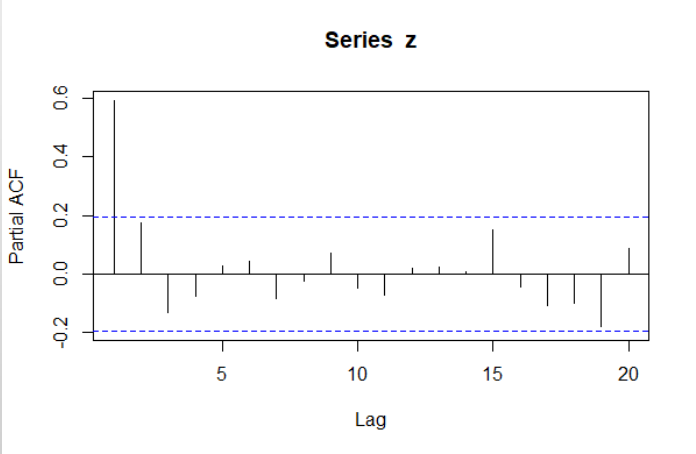
ex8_2b.txt

- 모형 인식 AR(1) 인식



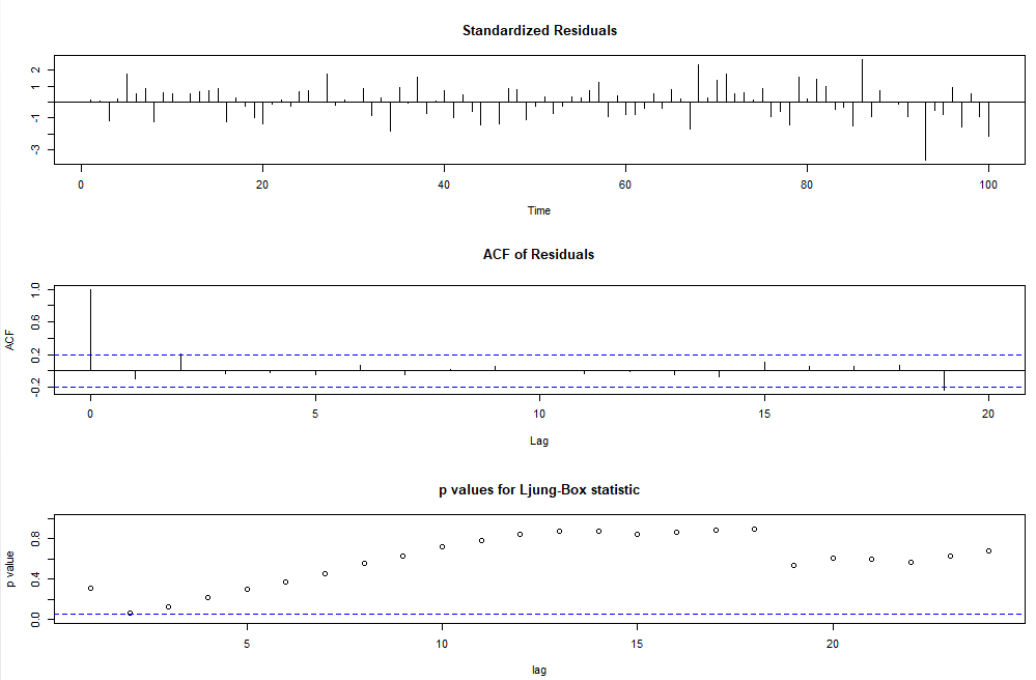


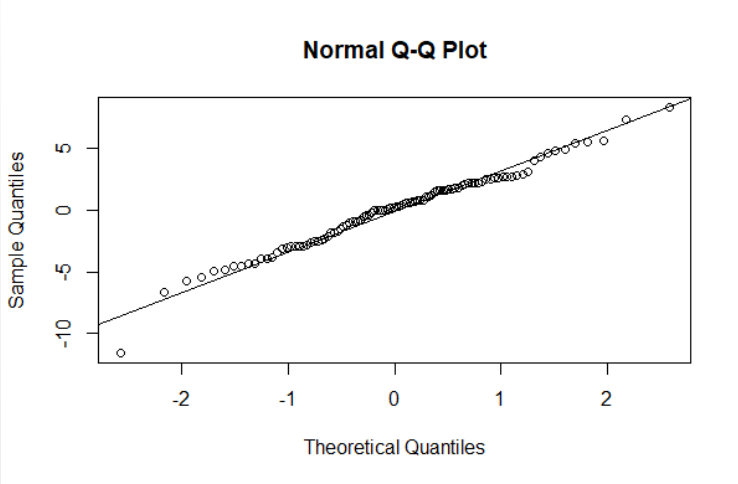


- 모형 추정

```
> fit <- arima(z,order=c(1,0,0))
> fit

Call:
arima(x = z, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.6231   100.4528
s.e.  0.0805     0.8253

sigma^2 estimated as 9.962:  log likelihood = -257.08,  aic =
520.16
```

- 모형 검진






```
> Box.test(fit$resid,lag=2,type="Ljung-Box",fitdf=1)

data:  fit$resid
X-squared = 5.6255, df = 1, p-value = 0.0177
```

- 함수 auto.arima( )에 의한 모형 인식

```
> library(forecast)
> auto.arima(z, stepwise=FALSE)
Series: z
ARIMA(1,0,2) with non-zero mean

Coefficients:
         ar1      ma1     ma2      mean
      0.5368  -0.0066  0.2947  100.4720
s.e.  0.1644   0.1651  0.1335    0.8385

sigma^2 estimated as 9.738:  log likelihood=-253.99
AIC=517.99   AICc=518.62   BIC=531.01
```

auto.arima( )에서 stepwise=FALSE 사용

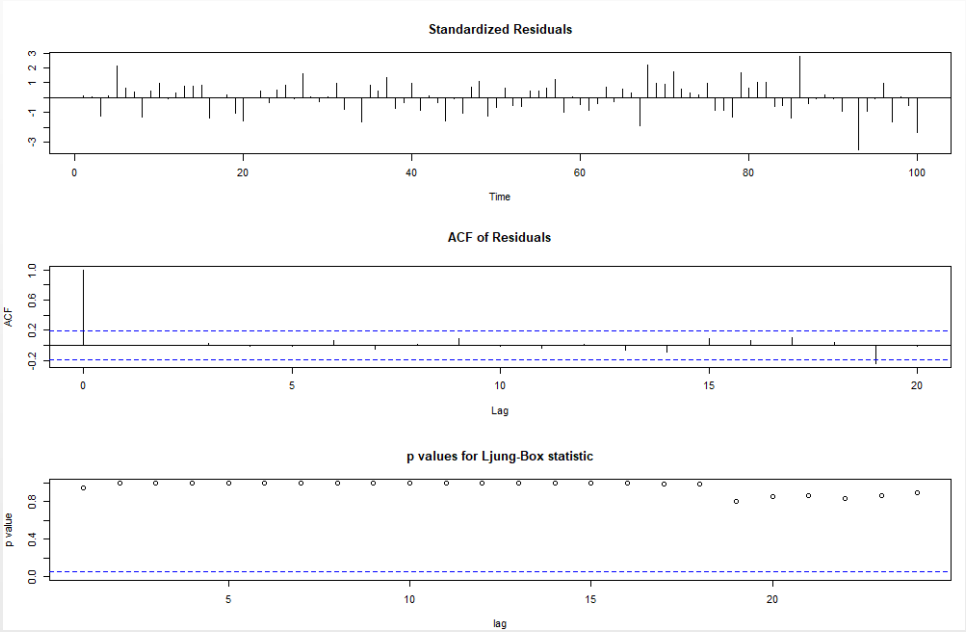
1) 모든 모형이 비교 대상
2) 속도가 느리다. 특히 계절형 모형의 경우

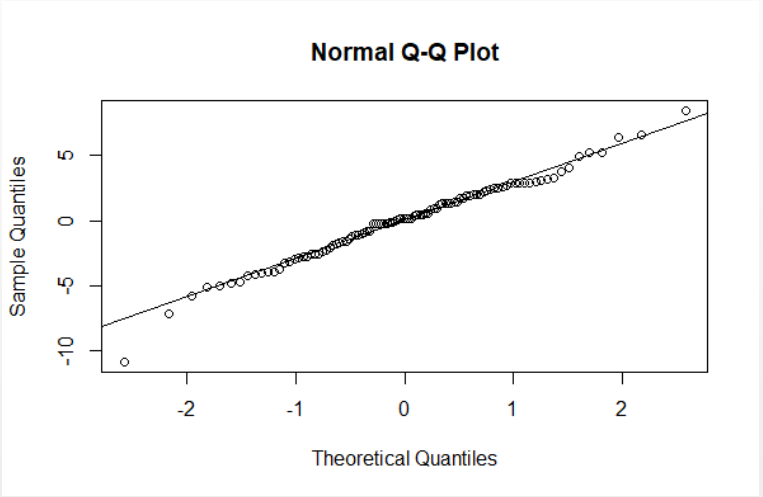
- 비유의적 모수 제외
- 다시 적합

- 비유의적인 모수 제외하고 모형 다시 적합

```
> fit1 <- arima(z,order=c(1,0,2),fixed=c(NA,0,NA,NA))
> fit1

Call:
arima(x = z, order = c(1, 0, 2), fixed = c(NA, 0, NA, NA))

Coefficients:
         ar1  ma1     ma2  intercept
      0.5315    0  0.2974   100.4733
s.e.  0.0961    0  0.1149     0.8344

sigma^2 estimated as 9.348:  log likelihood = -253.99,  aic
= 515.99
```

- 모형 검진


Standardized Residuals
Time
ACF of Residuals
ACF
Lag
p values for Ljung-Box statistic
p value
lag


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles

- ARMA(1,2) 모형의 과대적합

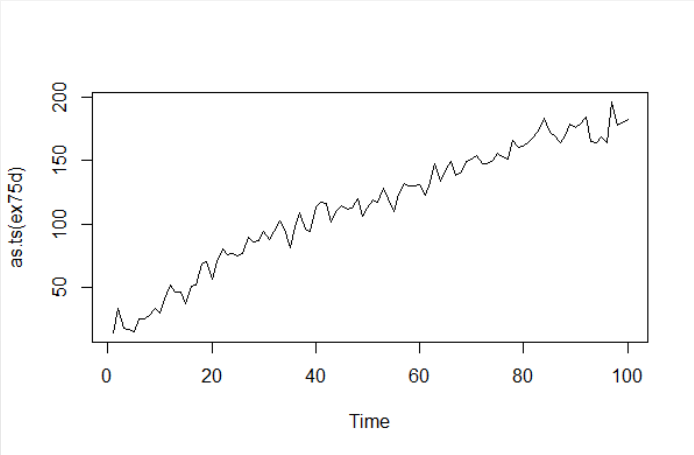
```
> confint(arima(z,order=c(2,0,2),fixed=c(NA,NA,0,NA,NA)))
                  2.5 %       97.5 %
ar1          0.33337949    0.7332027
ar2         -0.29609151    0.2801875
ma1                  NA           NA
ma2         -0.01892901    0.6266446
intercept 98.85024877 102.1018302

> confint(arima(z,order=c(1,0,3),fixed=c(NA,0,NA,NA,NA)))
                  2.5 %       97.5 %
ar1          0.33862742    0.7195490
ma1                  NA           NA
ma2          0.07469215    0.5200800
ma3         -0.17480342    0.2018783
intercept 98.82952467 102.1166314
```

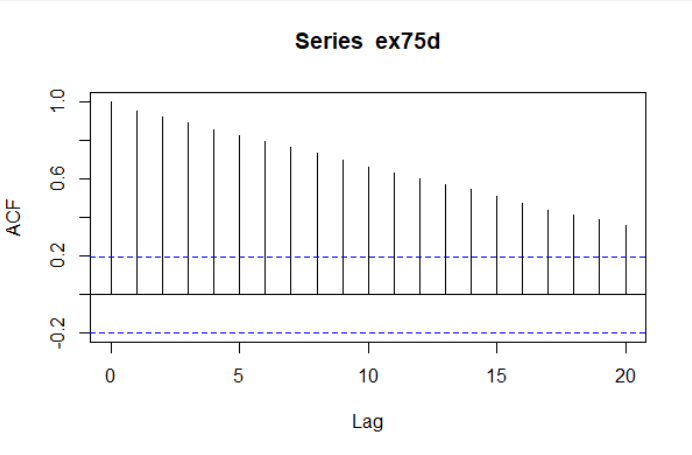
추가된 모수 모두 비유의적 → ARMA(1,2) 선정

연습문제: ex7_5d.txt

```
> ex75d <- scan("D:/Data/ex7_5d.txt")
Read 100 items
```

- 정상성 만족 확인

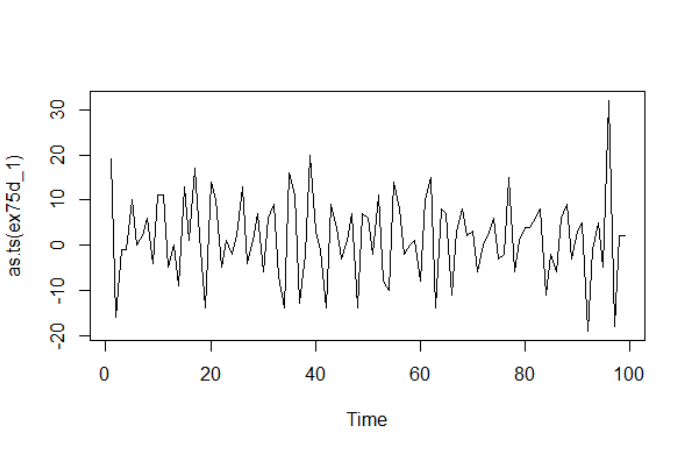





```
> library(forecast)
> ndiffs(ex75d)
[1] 1
```

1차 차분 필요

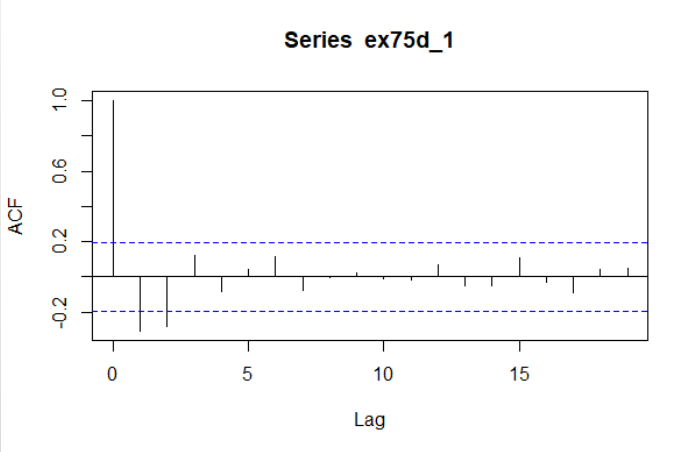
- 1차 차분 자료 정상성 확인 및 모형 인식

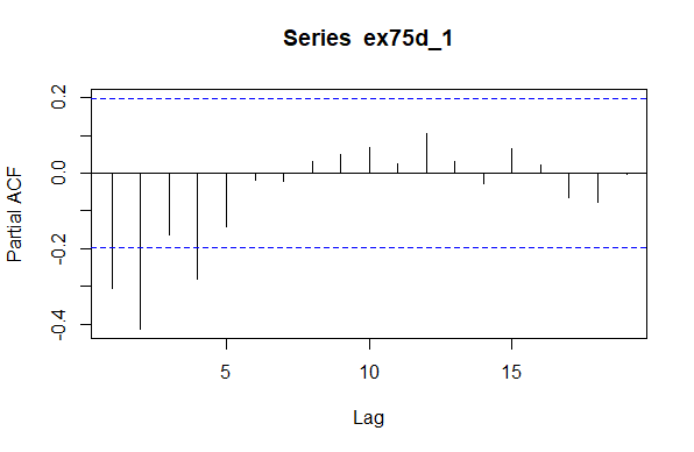


모형 인식

ARIMA(0,1,2)

ARIMA(1,1,1)
ARIMA(1,1,2)
ARIMA(2,1,1)
ARIMA(2,1,2)





1) ARIMA(0,1,2) 모형 적합

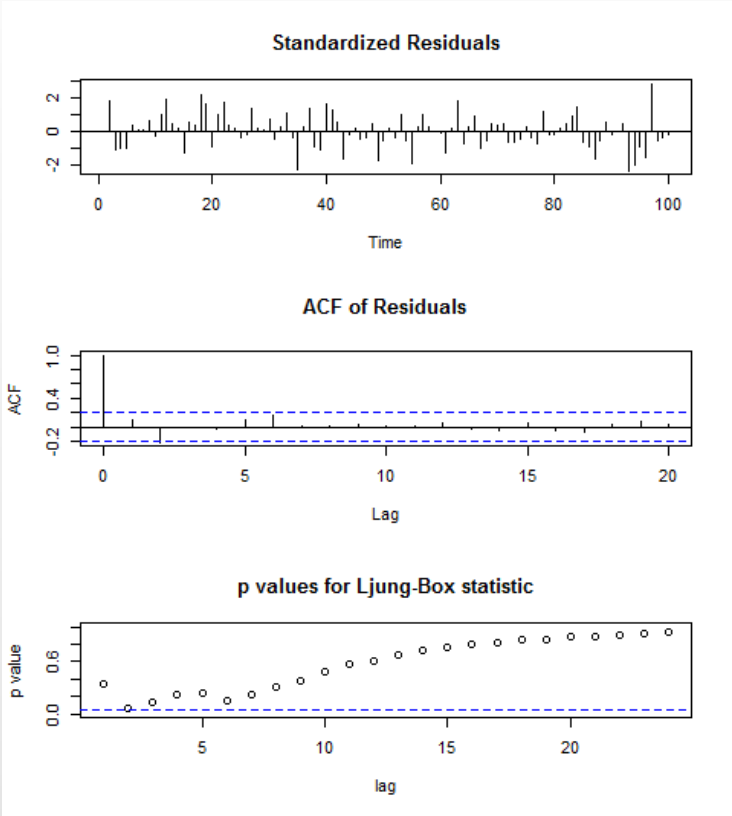
```
> library(forecast)
> Arima(ex75d, order=c(0,1,2), include.drift = TRUE)
Series: ex75d
ARIMA(0,1,2) with drift

Coefficients:
          ma1      ma2   drift
      -0.5276  -0.1896  1.6713
s.e.   0.1245   0.1207  0.2241

sigma^2 estimated as 59.88:  log likelihood=-341.84
AIC=691.67   AICc=692.1   BIC=702.05
```

ma2 모수 비유의적 → ARIMA(0,1,1) 모형

- ARIMA(0,1,1) 모형의 적합

```
> fit1 <- Arima(ex75d, order=c(0,1,1), include.drift = TRUE)
> fit1
Series: ex75d
ARIMA(0,1,1) with drift

Coefficients:
          ma1   drift
      -0.7025  1.6620
s.e.   0.0665  0.2363

sigma^2 estimated as 60.7:  log likelihood=-343.05
AIC=692.1   AICc=692.35   BIC=699.89
```

- ARIMA(0,1,1) 모형의 검진



- 부적합한 모형

- 과대적합 모형
  ARIMA(0,1,2) & ARIMA(1,1,1):

	2.5 %	97.5 %
ma1	-0.7716288	-0.28352060
ma2	-0.4260835	0.04691902
drift	1.2321333	2.11047020

	2.5 %	97.5 %
ar1	-0.1078336	0.3862985
ma1	-0.8935299	-0.6123492
drift	1.2201540	2.1136465

```
> Box.test(fit1$resid,lag=2,type="Ljung-Box",fitdf=1)$p.value
[1] 0.0179917
```

2) 혼합 모형의 적합: 최소 AIC, BIC 모형 선택 ARIMA(2,1,1) 모형

```
> Arima(ex75d, order=c(1,1,1),include.drift = TRUE)$bic
[1] 703.2355
> Arima(ex75d, order=c(1,1,2),include.drift = TRUE)$bic
[1] 703.3995
> Arima(ex75d, order=c(2,1,1),include.drift = TRUE)$bic
[1] 702.6889
> Arima(ex75d, order=c(2,1,2),include.drift = TRUE)$bic
[1] 706.7262
```

```
> Arima(ex75d, order=c(1,1,1),include.drift = TRUE)$aic
[1] 692.855
> Arima(ex75d, order=c(1,1,2),include.drift = TRUE)$aic
[1] 690.4239
> Arima(ex75d, order=c(2,1,1),include.drift = TRUE)$aic
[1] 689.7133
> Arima(ex75d, order=c(2,1,2),include.drift = TRUE)$aic
[1] 691.1555
```

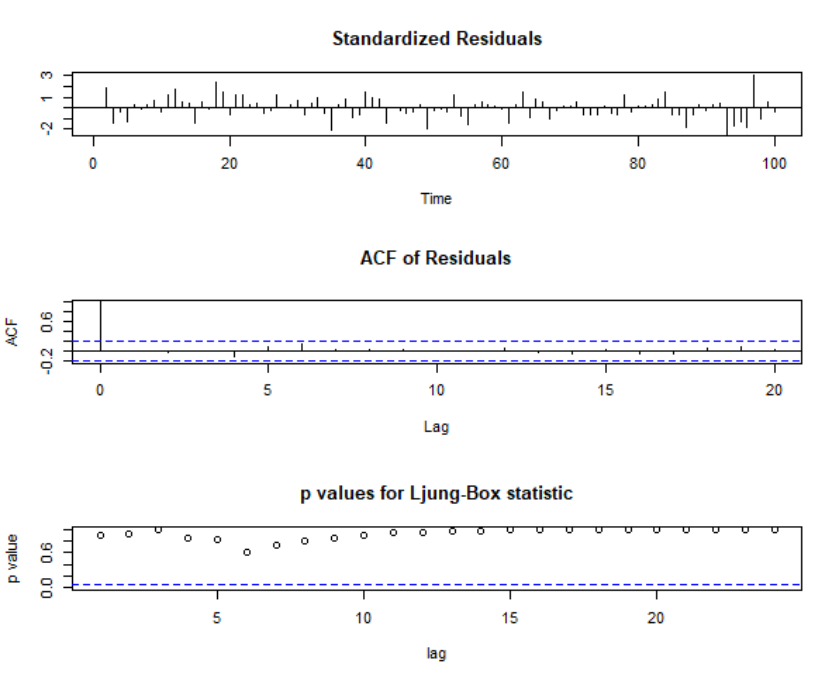
- ARIMA(2,1,1) 모형 적합

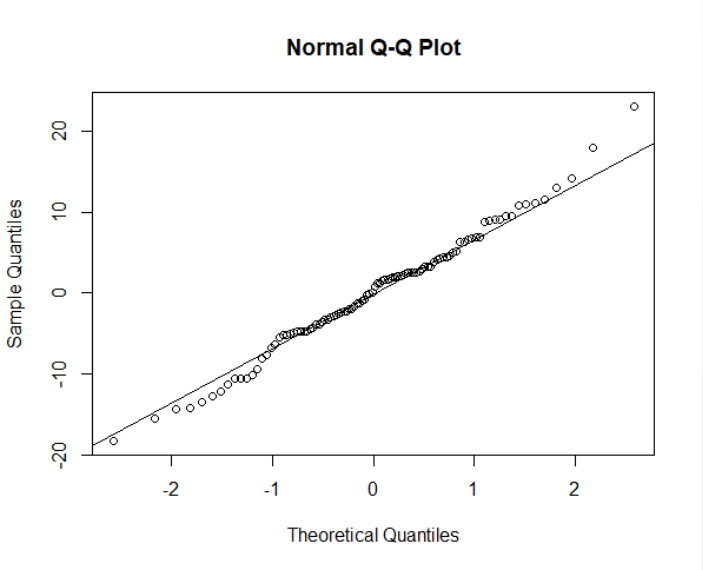
```
> fit2 <- Arima(ex75d, order=c(2,1,1), include.drift = TRUE)

> fit2
Series: ex75d
ARIMA(2,1,1) with drift

Coefficients:
          ar1      ar2      ma1   drift
      0.0412  -0.2746  -0.6231  1.6719
s.e.  0.1333   0.1141   0.1103  0.2342

sigma^2 estimated as 58.03:  log likelihood=-339.86
AIC=689.71   AICc=690.36   BIC=702.69

> confint(fit2)
            2.5 %      97.5 %
ar1    -0.2201178  0.30250218
ar2    -0.4981750 -0.05109239
ma1    -0.8393299 -0.40694244
drift   1.2129095  2.13080659
```

- ARIMA(2,1,1) 모형에서 ar1 모수 제거 후 다시 적합

```
> fit2.1 <- Arima(ex75d, order=c(2,1,1), include.drift = TRUE,
                  fixed=c(0,NA,NA,NA))

> fit2.1
Series: ex75d
ARIMA(2,1,1) with drift

Coefficients:
      ar1      ar2      ma1   drift
        0  -0.2860  -0.5988  1.6706
s.e.    0   0.1072   0.0849  0.2389

sigma^2 estimated as 58.09:  log likelihood=-339.9
AIC=687.81   AICc=688.23   BIC=698.19
```

- 모형 검진


Standardized Residuals
Time
ACF of Residuals
ACF
Lag
p values for Ljung-Box statistic
p value
lag


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles

- ARIMA(2,1,1) 모형에 대한 과대적합

```
> confint(Arima(ex75d, order=c(3,1,1), include.drift = TRUE,
                fixed=c(0,NA,NA,NA,NA)))
            2.5 %      97.5 %
ar1            NA          NA
ar2    -0.4978847 -0.07691048
ar3    -0.2146449  0.18012631
ma1    -0.7651125 -0.42556878
drift   1.2060652  2.13716849

> confint(Arima(ex75d, order=c(2,1,2), include.drift = TRUE,
                fixed=c(0,NA,NA,NA,NA)))
            2.5 %      97.5 %
ar1            NA          NA
ar2    -0.5095385 -0.02199123
ma1    -0.7832922 -0.37066219
ma2    -0.2623775  0.18707884
drift   1.2139814  2.12895114
```

추가된 모수 비유의적 → ARIMA(2,1,1) with $\phi_1 = 0$ 최종모형으로 사용 가능

3) 함수 auto.arima( )에 의한 모형 추천

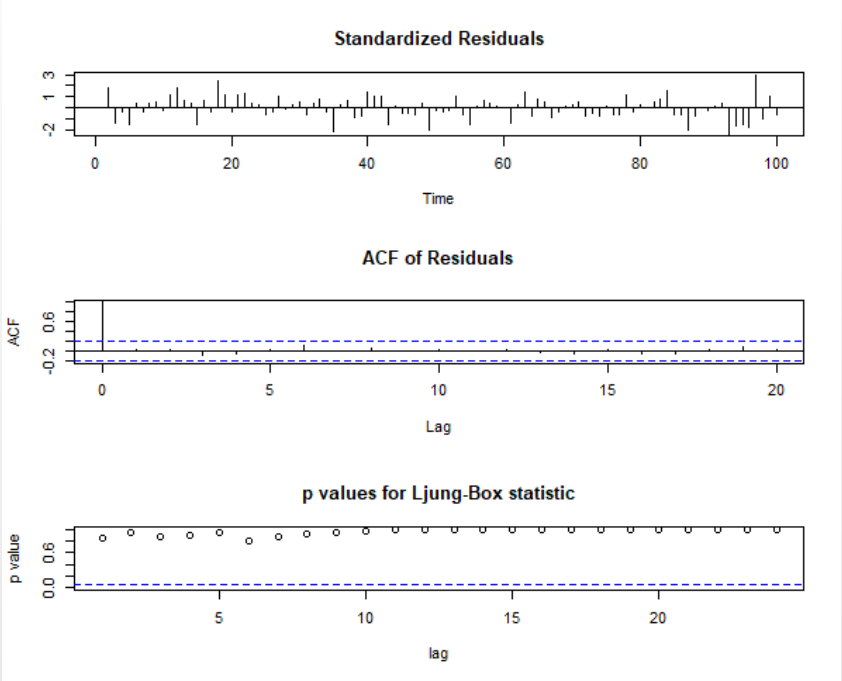
- 함수 auto.arima( )는 AIC를 최소로 하는 모형 탐색이 default
- BIC를 최소로 하는 모형을 탐색하고자 하는 경우에는
  auto.arima(ex75d, ic="bic")

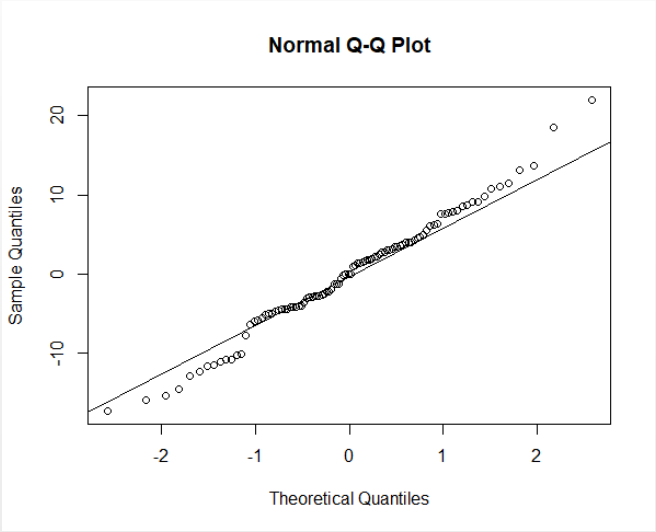
```
> auto.arima(ex75d,stepwise=FALSE)
ARIMA(0,1,3) with drift
Coefficients:
          ma1      ma2     ma3   drift
      -0.6022  -0.3651  0.2979  1.6658
s.e.   0.1045   0.1118  0.1172  0.2483

> auto.arima(ex75d,ic="bic",stepwise=FALSE)
ARIMA(0,1,1) with drift
Coefficients:
          ma1   drift
      -0.7025  1.6620
s.e.   0.0665  0.2363
```

- Arima(0,1,3) with drift 모형 적합

```
> fit3 <- Arima(ex75d,order=c(0,1,3),include.drift=TRUE)

> fit3
Series: ex75d
ARIMA(0,1,3) with drift

Coefficients:
          ma1      ma2     ma3   drift
      -0.6022  -0.3651  0.2979  1.6658
s.e.   0.1045   0.1118  0.1172  0.2483

sigma^2 estimated as 56.61:  log likelihood=-338.71
AIC=687.43   AICc=688.07   BIC=700.4

> confint(fit3)
             2.5 %     97.5 %
ma1   -0.80694420 -0.3974247
ma2   -0.58417343 -0.1460184
ma3    0.06824973  0.5276037
drift  1.17912663  2.1523963
```

- ARIMA(0,1,3) 모형 검진





- ARIMA(0,1,3) 모형 과대적합

```
> confint(Arima(ex75d,order=c(1,1,3),include.drift=TRUE))
              2.5 %     97.5 %
ar1     0.007084997  0.8608423
ma1    -1.443433477 -0.6033568
ma2    -0.553314783  0.2881898
ma3     0.170452848  0.6200290
drift   1.062203337  2.2736501

> confint(Arima(ex75d,order=c(0,1,4),include.drift=TRUE))
             2.5 %     97.5 %
ma1    -0.79629657 -0.4023714
ma2    -0.65444266 -0.1494722
ma3     0.02785128  0.4992412
ma4    -0.11249099  0.3089903
drift   1.14473771  2.1930321
```

ARIMA(1,1,3)

추가된 모수 유의적/ ma2 비유의적
→ 추가 분석

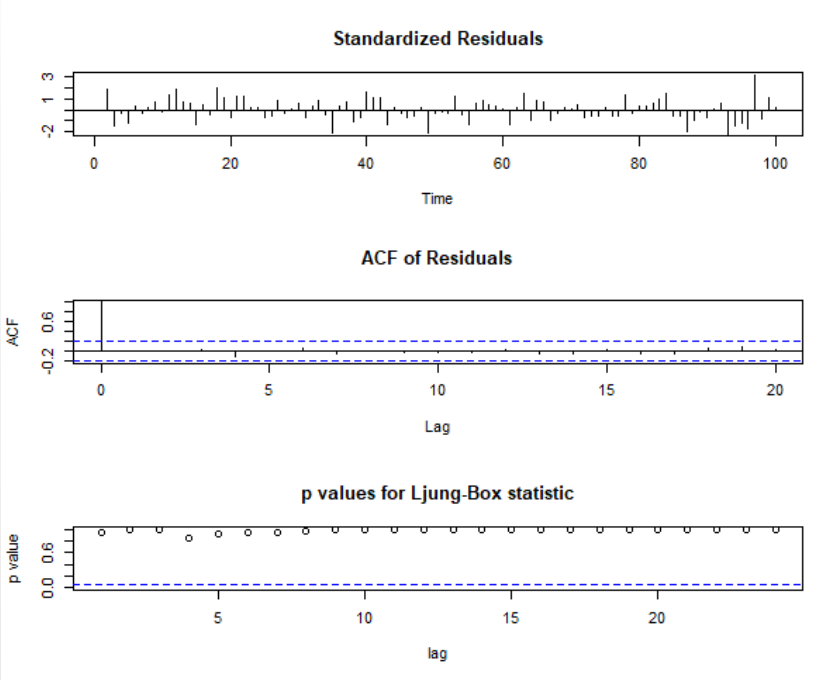
ARIMA(0,1,4)
추가된 모수 비유의적

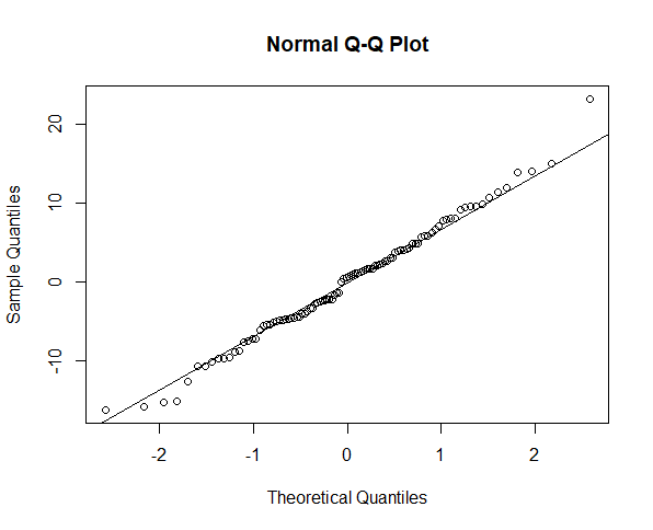
- ARIMA(1,1,3) with ma2=0 모형 적합

```
> fit4 <- Arima(ex75d,order=c(1,1,3),include.drift=TRUE,
                fixed=c(NA,NA,0,NA,NA))

> fit4
Series: ex75d
ARIMA(1,1,3) with drift

Coefficients:
         ar1      ma1  ma2     ma3   drift
      0.5180  -1.1324    0  0.3466  1.6664
s.e.  0.1385   0.0845    0  0.0782  0.3247

sigma^2 estimated as 56.27:  log likelihood=-337.99
AIC=685.97   AICc=686.62   BIC=698.95
```

ARIMA(0,1,3) 모형보다 작은 값의 AIC, AICc, BIC

- ARIMA(1,1,3) with ma2=0 모형 검진





- ARIMA(1,1,3) with ma2=0 모형에 대한 과대적합

```
> confint(Arima(ex75d,order=c(2,1,3),include.drift=TRUE,
                fixed=c(NA,NA,NA,0,NA,NA)))
            2.5 %     97.5 %
ar1     0.2300082  0.8010240
ar2    -0.2982268  0.2019625
ma1    -1.3220668 -0.9064129
ma2            NA         NA
ma3     0.1765349  0.5109311
drift   1.0520675  2.2877721

> confint(Arima(ex75d,order=c(1,1,4),include.drift=TRUE,
                fixed=c(NA,NA,0,NA,NA,NA)))
             2.5 %     97.5 %
ar1     0.20357686  0.9750608
ma1    -1.48257305 -0.9050508
ma2             NA         NA
ma3     0.03500506  0.8576986
ma4    -0.35740462  0.2078702
drift   1.04260684  2.2813648
```

추가된 모수 비유의적 → ARIMA(1,1,3) with ma2=0 최종모형 선택 가능

4) 최종 모형 선택

```
> fit2.1
Series: ex75d
ARIMA(2,1,1) with drift

Coefficients:
      ar1      ar2      ma1   drift
        0  -0.2860  -0.5988  1.6706
s.e.    0   0.1072   0.0849  0.2389

sigma^2 estimated as 58.09:  log likelihood=-339.9
AIC=687.81   AICc=688.23   BIC=698.19

> fit4
Series: ex75d
ARIMA(1,1,3) with drift

Coefficients:
         ar1      ma1  ma2      ma3   drift
      0.5180  -1.1324    0   0.3466  1.6664
s.e.  0.1385   0.0845    0   0.0782  0.3247

sigma^2 estimated as 56.27:  log likelihood=-337.99
AIC=685.97   AICc=686.62   BIC=698.95
```

- 두 모형 모두 사용 가능
- AIC, BIC 등을 비교하여 최종 모형 선택

ARIMA(1,1,3) 최종 모형으로 선택

- 최종 모형식

```
> fit4$coef
       ar1         ma1         ma2         ma3       drift
 0.5179900 -1.1324205   0.0000000   0.3466253   1.6663607
```

$$(1-0.518B)(1-B)Z_t = \delta + (1-1.1324B+0.3466B^3)\varepsilon_t$$

$$\delta = \mu(1-\phi) = 1.6664 \times (1-0.518)$$
$$= 0.8032$$